Interlanguage Phonology and the Pronunciation of English Final Consonant Clusters by Native Speakers of Vietnamese

Ngan Nguyen
Ohio University

The pronunciation of English syllable codas by second language (L2) learners of English, especially those whose native languages (L1s) do not have phonetic features similar to English, has received much attention in L2 research. Vietnamese, for example, does not have consonant clusters which are allowed in English in various word-positions. Vietnamese L2 learners of English have been found to have problems with pronouncing English consonant clusters, particularly those in word-final positions. The present research aimed to yield more insights into the interlanguage phonology of Vietnamese learners of English in this regard. Specifically, it was designed to identify (1) final clusters that Vietnamese L2 learners of English had difficulty producing accurately, (2) phonological processes that appeared to be at work when a final cluster was pronounced inaccurately, and (3) if task type had an effect on accuracy in final consonant production. The participants in this study were five Vietnamese students, enrolled in a U.S. university, whose performance on three tasks— reading a wordlist, reading a text, and semi-structured interview—was audio recorded and transcribed. The results showed that final consonant clusters which consisted of voiced obstruents (e.g., /bd/, /vz/) were more difficult than those containing voiceless obstruents. Final consonant clusters containing a liquid (e.g., /rt/, /lθ/) were also more problematic than those containing a nasal. Deletion was found to be the most common modification strategy and task type was found to have an effect on the articulation of final consonant clusters.

Whereas much research with a focus on syllable codas has been done on English as a second language (ESL), native speakers (NSs) of Portuguese, Arabic, and Chinese, not many researchers have examined NSs of Vietnamese in this regard. Not until the 1980s, when there was an influx of Vietnamese immigrants to the United States after the Vietnam War, did researchers such as Sato (1984, 1985) and Benson (1986) begin to investigate the L2 phonology of Vietnamese immigrants upon noticing their erroneous production of syllable onsets and codas. Findings from these studies have confirmed that consonant clusters are one area of difficulty Vietnamese ESL learners have. In this regard, Hwa-Froelich, Hodson, and Edwards (2002) did a rather comprehensive investigation into Vietnamese phonology, comparing and contrasting its phonological characteristics with English, in order to see how these features may be carried by

Vietnamese learners into their L2 English phonology. They also suggested that Vietnamese learners may have problems with both suprasegmental and segmental features in English, particularly word final consonants. Nevertheless, little research has been done on the acquisition of final consonant clusters by learners from this particular language background; thus, the present study aims at gaining more insights into the interlanguage phonology of Vietnamese learners of English. In particular, it focuses on the production of two member final consonants (2MFCs) by advanced Vietnamese learners of English and the common modification strategies these learners employ in their pronunciation of final consonant clusters. Task type is also examined as a potential factor that may have effects on the pronunciation output.

The research focuses on the following questions:

(1) What types of 2MFCs are the most difficult for Vietnamese ESL learners to pronounce?

(2) What are some of the most common modification strategies employed by Vietnamese ESL learners in their production of 2MFCs?

(3) Do modification strategies of 2MFCs vary across tasks?

Markedness, Linguistic Environment, and Task Variation

Second language phonology is unique in that L2 learners who start learning another language after puberty, do not usually achieve nativelike L2 phonology, while they may do quite well in other language areas such as syntax, morphology, or discourse (Tarone, 1987). Scovel (1969) termed this phenomenon the "Joseph Conrad phenomenon" to commemorate the British Polish-born novelist who did not learn to speak English until he was in his 20s and, even though he acquired perfectly English morpho-syntax, he spoke English with a strong Polish accent all his life. It used to be believed that interlanguage phonology was simply a result of direct transfer from one's L1 phonology; therefore, it was assumed that research in this area could hardly bring in any new discoveries (Ioup &Weinberger, 1987). Although much research has suggested that L1 phonetics and phonology have a tremendous influence on L2 pronunciation (Odlin, 1989), there is evidence supporting the notion that L1 is no longer the single source of all phonological errors (Hansen, 2001; Ioup &Weinberger, 1987; Odlin, 1989; Osburne, 1996). According to Major's Ontogeny Model (1987), L1 influence seems to have strong effects in the early stages of L2 acquisition; however, later on, as learners' proficiency increases, developmental factors are likely to become more prominent. Additionally, universal preferences, markedness, sociolinguistic and psychological factors have also been found to influence interlanguage phonology (Bebee, 1987; Broselow, 1987; Eckman, 1977; Greenberg, 1965, Sato, 1984, 1985; Tarone, 1987; Weinberger, 1987).

Generally, research on markedness effects has yielded some useful predictions of problematic areas for different language learners. The Markedness Differential Hypothesis (Eckman, 1977) predicts that the areas of the target language which differ from and are more marked than the L1 will be difficult for learners to acquire. In light of this hypothesis, Eckman (1977) suggested developing markedness hierarchies to predict interlanguage variations. One of the predictions of the hypothesis is that voiced contrast is the most marked in final positions and least marked in initial positions. In this regard, Benson (1986) inspected the pronunciation of consonant clusters of Vietnamese L2

learners of English, based on Greenberg's (1965) universal generalizations of consonant sequences and found the following hierarchies: (1) for final consonant clusters containing a nasal, voiceless homorganic nasal obstruent clusters are less marked than voiced heterorganic nasal obstruent clusters (e.g., /nt/, alveolar nasal + voiceless alveolar stop, is less marked than /md/, bilabial nasal + voiced alveolar stop); (2) for clusters containing obstruents only, voiceless obstruent clusters are less marked than voiced obstruent clusters (e.g., /kt/, voiceless velar stop + voiceless alveolar stop, is less marked than /bd/, voiced bilabial stop + voiced alveolar stop). Further research (Benson, 1988; Osburne, 1996; Sato, 1984) has supported Benson's hierarchies.

Markedness in consonant clusters has received a lot of research attention based on the observation that they do not occur frequently across languages. There is little doubt now that final consonant clusters are more marked than the ones in initial or medial position (Anderson, 1987; Benson, 1986; Major, 1996; Sato, 1984; Weinberger, 1987). The languages investigated in most studies have been ones that either allow no consonant clusters in any position (e.g., Mandarin and Vietnamese) or allow certain consonant clusters in word final position only (e.g., Arabic), or word initial position only (e.g., Portuguese). One common finding across these studies has been that NSs of these languages show great difficulty with acquiring English final consonant clusters but not quite so many with the ones in initial position. Moreover, learners from different language backgrounds seem to have different preferences for cluster simplification strategies, ranging from vowel epenthesis, to substitution or deletion of certain segments in final consonant clusters. While the L1 Mandarin and Arabic speakers seem to prefer deletion (Anderson, 1987), the Portuguese natives' most common strategy is a substitution of the voicing contrast (Major, 1996). In addition, Weinberger (1987) noticed that cluster simplification strategies can vary across level of proficiency. Similarly, Hansen (2001) had a group of intermediate learners with whom she found a high rate of epenthesis for two-member final clusters (codas), and a tendency to eliminate one of the members in codas consisting of three members.

In the case of Vietnamese ESL learners, Sato (1984) did not find epenthesis to be a common strategy used in the pronunciation of final consonant clusters while Benson (1986) found a considerable preference for insertion of schwa after obstruents, suggesting an optimal regression to the basic consonant-vowel (CV) structure (/kt/ → /ktə/). This finding corresponds to Tarone's (1979) finding that learners fall back to the simple universal CV structure, even though the syllable sequence is allowed in their L1. As to strategies, both Benson and Sato observed devoicing, deletion, and substitution as common cluster simplification strategies. These findings were confirmed by Osburne's (1996) case study of a male advanced speaker subject who still showed various forms of interlanguage phonology. For example, in his speech, he tended to omit final consonants which were not allowed in word final position in Vietnamese. Also, while he rarely employed complete deletion, Osburne (1996) noticed occasional dropping of inflectional morphemes which she termed 'short forms' of plural or third person singular markers (words such as *resources*, *charges*, *expenses* are examples of 'long forms'). Osburne, however, was cautious about categorizing those as forms of interlanguage phonology since they might have well been a result of grammar errors (e.g., *he talk** vs. *he talked*). This phenomenon of omitting final inflectional morphemes after a consonant was also noted by Honey (1987). In fact, NSs of English also simplify final clusters; however, the consonants to be eliminated should never be, by rule, the inflectional morphemes (Celce-Murcia, Brinton, & Goodwin, 1996).

Studies focusing on the linguistic environment have found that variation in the pronunciation of consonant clusters may also be influenced by the immediately preceding or following segments. Vocalic and consonantal environments as well as pausing were all found to interfere with cluster modification (Baley, 1996; Benson, 1988; Hansen, 2001; Tarone, 1987). Vietnamese ESL learners have been shown to usually not pronounce final consonantal segments after diphthongs since most diphthongs in Vietnamese appear in open syllables (Benson, 1988; Osburne, 1996; Sato, 1984). However, I should note here again that interlanguage phonology is not influenced by L1 syllable structure alone. As mentioned earlier, one of the factors causing errors could be learners' preference for universal basic CV structures (Sato, 1984; Tarone, 1979). Osburne (1996), for instance, noticed that her subject's native dialect had caused him to acquire an *r*-less coda in his English speech even though Vietnamese does not allow the alveolar approximant in syllable codas.

Stress also seem to play an important role in L2 phonology. Cutler and Noris (1988), for example, examined the role of strong syllables in lexical access and noted a consistent accuracy in recognition of forms which are stressed. Labov (1989) further suggested that syllable stress is a constraint on cluster modification. Also, Barley (1996) found that Chinese learners of English exhibited cluster reduction more on unstressed clusters than on stressed ones.

Research in L2 acquisition has also investigated whether task variation has an influence on interlanguage phonology. Labov (1971) noted that different forms are likely to occur depending on the speech situation. Tarone (1979, 1982, 1983) took on the idea and proposed the Chameleon model with five axioms: style-shifting across social situations and topics, attention to speech, vernacular, formality, and good data. Her claim was that there may be more variations when less attention is paid to the form of the language since speakers in this case might focus more on content, or be carried away with anxiety or relaxation. Attention increases or decreases depending on context formality, which is associated with a continuum of styles. For example, reading a list of minimal pairs should be considered the most formal style since most attention is paid to speech and a higher degree of accuracy should be expected, whereas casual speech should be considered a much less formal situation where error is likely to override correctness (Dickerson & Dickerson, 1977). Between the extremes of the continuum, however, there are immediate styles such as careful speech, paragraph reading, and wordlists (Labov, 1971).

There is some research, though, that has not quite confirmed Tarone's Chameleon Model and raised doubt if attention to speech could be an adequate explanation of variation. For instance, the data collected by Sato (1985) across three tasks (imitation, oral reading, and conversation) showed that form variations were unpredictable, which lent support to Gass's (1980) suggestion that there are more processes going on when learners perform different tasks than just attention to speech alone. These processes can involve decoding and accepting data (e.g., an acceptability judgment task), or using knowledge of the language to judge and/or produce a certain form (e.g., a sentence combining task).

In summary, the findings regarding the acquisition of final consonant clusters by L2 learners suggest that learners' errors can be predicted by areas of markedness, L1 transfer, linguistic environment, and task variation. The manifestation of these factors also depends on learners' language proficiency, with learners from the lower proficiency

levels having the most inaccurate production rate. It should also be noted that highly advanced learners still exhibit erroneous production, which is nothing new but an idée fixe as far as interlanguage phonology is concerned. In addition, learners from different L1 background appear to have different preferences for modification strategies. Vietnamese ESL learners, in particular, have been observed to employ some common modification strategies such as deletion, substitution, epenthesis, or a change in features of one or all members of the clusters. As far as variation across tasks is concerned, studies have not all agreed that errors might occur more often as the informality of tasks increases.

Based on the theoretical background discussed above regarding markedness, linguistic environment, task variation and the findings of studies focusing particularly on the pronunciation of English consonant clusters by Vietnamese L2 learners of English (Benson, 1988; Osburne, 1996; Sato, 1984, 1985), this research puts forth the following hypotheses:

(1) Final consonant clusters consisting of voiced stops, fricatives, or liquids will be the most difficult for Vietnamese L2 learners of English to pronounce.

(2) Eliminating marked segmental members of the clusters and devoicing will be prominent modification strategies for Vietnamese L2 learners of English to employ when dealing with final consonant clusters in English.

(3) Vietnamese L2 learners of English will make more pronunciation errors of final consonant clusters in the interview task than the wordlist reading task.

The final clusters which will be considered in this study combine two consonants as described below:

(1) Clusters containing obstruents only:

SS: stop stop (e.g., /pt/, /bd/)

FF: fricative fricative (e.g., /fs/, /vz/)

SF: stop fricative (e.g., /ps/, /dz/)

FS: fricative stop (e.g., /st/, /zd/)

(2) Clusters containing a nasal:

NS: nasal stop (e.g., /nk/)

NF: nasal fricative (e.g., /nz/)

(3) Clusters containing a liquid (/l/ or /ɹ/):

LS: liquid stop (e.g., /lt/, /rb/)

LF: liquid fricative (e.g.: /lv/, /rz/)

In terms of modification strategies, the study will aim at finding common patterns of strategies, such as devoicing, segmental substitution, segmental assimilation, epenthesis, cluster reduction or deletion, in an attempt to provide a possible explanation of those preferences.

The study

Participants

The participants of this study were five Vietnamese ESL speakers (three males and two females), age 20 to 31. They were students who pursued academic degrees in different areas of study at one American university. Their paper-based TOEFL scores ranged from 550 to 575, which range is usually considered to indicate an advanced level of proficiency according to the Proficiency Guidelines of the American Council for the Teaching of Foreign Languages (ACTFL Proficiency Guidelines). As a requirement for their teaching or graduate assistantships, the participants have taken and passed the institutional Speak Test.

All informants participating in the study had started learning English after the age of twelve, which means after the Critical Period (Lenneberg, 1967). At the time of the data collection, the participants have studied English for ten years on average (ranging from five to sixteen years) and had stayed in the United States for on average eleven months (ranging from six to thirteen months).

Instruments

Data were collected via three tasks: a wordlist, reading text, and semi-structured interview.

*Wordlist (*see Appendix A)

The wordlist contained 72 words which selection was based on different types of 2MFCs. More specifically: (1) Clusters that contained a nasal; (2) Clusters that contained a liquid; (3) Clusters with voiced obstruents only; (4) Clusters with voiceless obstruents only.

The words were presented in sets of three, where the words in the same set had the same ending cluster sounds. Within a set, the first two words formed a minimal pair and the third word was a two-syllable word, with the stress falling on the first syllable. The wordlist also contained some words which did not follow the rules mentioned above. These words (about 17% of the wordlist) either ended with three consonants (e.g., /rlz/) or with two consonants of a liquid and a nasal sound (e.g., /lm/).

In terms of formality of speech, the wordlist was considered the most formal among the three tasks and it was expected that the participants will pay great attention to the production of each single sound, thus, yielding the most correct pronunciation.

*Reading text (*see Appendix B)

The text was created by the researcher in consideration of the criteria applied to the creation of the wordlist. The text was designed to consist of items representative of four groups of sounds—i.e., voiceless obstruents, voiced obstruents, nasals, and liquids. The text was 552 words long, but data were collected only from the middle paragraph which contained 38 targeted words (i.e., words which ended with two member consonant clusters). The middle paragraph was chosen as it was assumed that the participants may become more relaxed after having read the first paragraph, so their pronunciation would be more natural when reading the second paragraph. The text reading task was considered to be less formal than the wordlist task.

Interview (see Appendix C)

The interview was semi-structured, consisting of casual questions about the participant's demographic background and a map description task. The participants were asked to describe a map which was specifically designed with street names and places to reflect patterns of focused codas as described for the wordlist. In this task, I had informal conversations with the participants for a few minutes before asking the interview questions. The data collected from this task was considered to be more natural than the data from the two tasks discussed above.

Table 1 summarizes the data collected from the five informants via the three tasks:

Table 1

Types of Final Consonant Clusters in the Study per Participant

Types of clusters	Wordlist	Text Reading	Interview
Voiceless obstruents	12	5	10
Voiced obstruents	12	5	6
Clusters containing a nasal	12	15	13
Clusters containing a liquid	24	11	14

Procedure

Each participant was asked to read the wordlist and the text while being audio-recorded. The participants were allowed to scan the wordlist before reading it. If there was a word that they were not sure how to pronounce or did not know, I pronounced it for them as a model. For the text, the participants were not presented with the material beforehand but were asked to read it with their normal reading speed at the time of the recording. After a participant finished the reading task, he/she was interviewed for about five to ten minutes in English.

Two volunteer NSs of English transcribed the recordings separately, using the IPA transcription symbols. The transcribers were two second year graduate students of linguistics who had formal training in linguistics and phonetics. The purposes of the study were communicated thoroughly to them in order to ensure accuracy and consistency of the data transcription. An inter-transcriber agreement of 77.6 % was obtained, which was within acceptable values. Next, the transcribed data were categorized as (1) expected production (with both consonants present in final position) and (2) modification. The modification category then was subcategorized into (a) deletion (of the first consonant in the cluster, the second consonant in the cluster, or both consonants), (b) epenthesis (of a vowel), and (c) feature change (voicing vs. devoicing, aspiration, substitution or insertion of another consonant).

Results

Figure 1 illustrates the participants' overall production, categorized by expected production and modification. It is clear that the percentage of expected production was very low, especially in the interview task (7.4%). There was also variation across the three tasks, with the highest percentage of expected production evident in the wordlist (19.7%), immediately followed by the text reading task (14.4%).

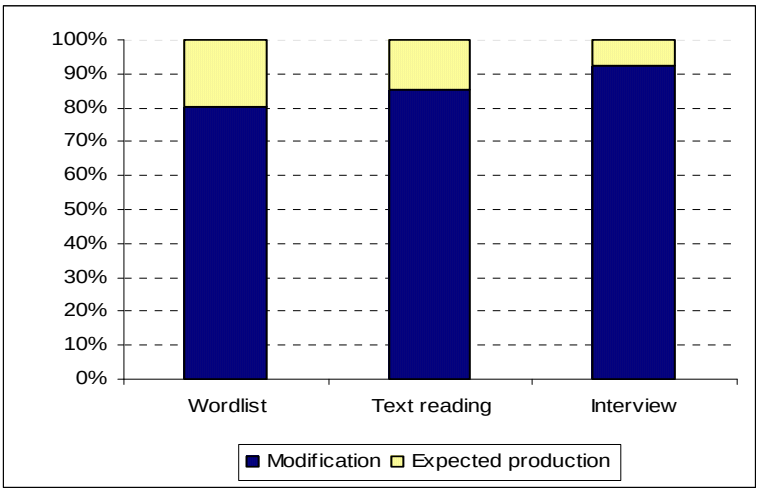

Figure 1: Overall results of participants' pronunciation of final consonant clusters.

Table 2 presents a more detailed description of the modification strategies employed by the participants, where feature change seemed to outnumber deletion and epenthesis in the wordlist task. However, in the text reading and interview tasks, deletion took the lead. There was no epenthesis found in the text reading task.

Table 2

Summary of Participants' Pronunciation of English Final Consonant Clusters across Tasks

Task/Production results	Wordlist	Reading text	Interview
Expected production	59	26	15
Modification:			
Deletion	99	112	132
Epenthesis	9		4
Feature change	133	42	51
Total number of tokens	300	180	202

A closer look at deletion as a modification strategy revealed that the deletion of the second consonant of a cluster was more common than that of the first or both. As can be seen in Table 3, the informants dropped both consonants in all three tasks with the highest frequency in the interview, where less attention to form was expected.

Table 3

Distribution of Deletion Types of Participants' Pronunciation of English Final Consonant Clusters across Tasks

Type	Wordlist (%)	Reading text (%)	Interview (%)
Deletion of the 1st consonant of the cluster	30.3	17.9	15.2
Deletion of the 2nd consonant of the cluster	57.6	58.9	46.2
Deletion of both consonants	11.1	23.2	38.6

A frequency count of voicing versus devoicing was performed and, not surprisingly, it was found that devoicing of clusters that ended with a voiced consonant (which usually makes the clusters voiced) was more common than voicing of clusters that ended with a voiceless consonant. It is also worth noting that the informants of this study showed a higher percentage (55%) of devoicing of voiced clusters in the wordlist task, where they were expected to be more conscious of the expected pronunciation. Notably, the cluster of a voiced stop and a voiced fricative (e.g., /bz/, /dz/) appeared to cause most problems, followed by a voiced fricative and a voiced stop, two voiced fricatives, liquid and voiced stop, and liquid and voiced fricative (e.g., /vd/, /vz/, /ld/, /lz/). In other words, clusters that ended with voiced obstruents were devoiced.

I would like to point out here that I did not have an ideal balance among all cluster types across the three tasks. The percentage of clusters ending with a voiced consonant in the reading text was 75%, compared to 25% of clusters ending with voiceless consonants. There was a more equal balance between these two types of consonant clusters in the wordlist (50% vs. 50%) and the interview (51.2% vs. 48.8%). Yet, the comparison was interesting because the informants might have been fully aware of the voiced characteristics of some clusters even though they did not produce them correctly. At the same time, the lower percentage of desired pronunciation in the text reading and interview may not mean that these informants produced more desired pronunciation of final consonant clusters when they were more relaxed or concerned with, let's say, content rather than linguistic form. It should also be noted that the voicing of clusters that ended with a voiceless consonant was found only with the wordlist but not with the other two tasks, which suggested that in more natural speech, voicing of clusters ending with voiceless sounds was not what these informants were going to do.

Table 4 displays the results for the expected production for each type of combination of final consonant clusters. It was difficult to find a solid pattern across all three tasks. However, it could be seen that combinations such as voiced stop + voiced stop (SvdSvd), voiced stop + voiced fricative (SvdFvd), voiced fricative + voiced stop (FvdSvd), voiced fricative + voiced fricative (FvdFvd), and liquid + voiced fricative (LFvd) were truly difficult for the informants of this study since none of them achieved more than 10% of expected production. For the combinations with a nasal sound, it was evident that they did not seem to cause any pronunciation problems for the participants. compared to the clusters of obstruents only or those that contained a liquid. Also, clusters that ended in a voiced fricative (SvdFvd, FvdFvd, or LFvd) were with a very low expected production; however, when a voiced fricative was preceded by a nasal, the production result seemed to be improved.

Table 4

Expected Production of Participants' Specific Combinations of Final Consonant Clusters

Types of clusters	Wordlist (%)	Reading text (%)	Interview (%)
SvlSvl	13.3	—	—
SvlFvl	40.0	40.0	13.3
FvlSvl	20.0	—	—
FvlFvl	33.3		
SvdSvd	—		
SvdFvd	—	—	—
FvdSvd	6.7	10.0	
FvdFvd	—		—
NSvl	66.7	50.0	20.0
NSvd	13.3	11.4	5.0
NFvl	13.3	60.0	40.0
NFvd	40.0	40.0	25.0
LSvl	46.7		—
LSvd	13.3	—	
LFvl	23.3	—	20.0
LFvd	3.3	5.0	—

As far as the obtained productions were concerned, final clusters appeared to be commonly reduced to a single voiceless fricative or voiceless stop. Final clusters that

contained only voiced obstruents induced the lowest expected production. Final clusters containing a nasal seemed to be produced with the nasal only. Combinations of a nasal and a voiced stop were found to be the most difficult since only seven out of seventy clusters of this type were produced as expected. The same pattern could be observed with final clusters containing a liquid.

Discussion

Consonant cluster reduction has been found to be common not only in L2 but also in L1 phonology. For instance, Wolfram (1998) observed a high frequency of reduction of consonant clusters across different English dialects with reduction of clusters in pre-consonantal position being more common than that in pre-vocalic position. According to Wolfram (1998), Vietnamese English has the highest percentage of consonant cluster reduction in pre-consonantal position, and only ranks after Native American Puebloan English for pre-vocalic consonant cluster reduction.

The present study focused on final consonant clusters that consisted of two consonants only. The overall results revealed that Vietnamese ESL speakers used modification strategies quite frequently. Discussions concerning the specific research questions are presented below.

The first research question asked what types of 2MFCs were the most difficult for Vietnamese ESL learners to pronounce. For 2MFCs of obstruents only, the voiceless clusters generated higher percentages of expected production than the voiced ones. The combination of a nasal and a voiceless obstruent also showed a higher frequency of desirable production, which held true for 2MFCs that contained a liquid too. It seemed that markedness played a role in determining the level of difficulty of these clusters since universally, voiced obstruents are more marked than voiceless (Greenberg, 1965). There were exceptions, however, as found in the case of NFvd clusters, which showed a comparatively high percentage of expected pronunciation (25% across the three tasks). In this particular case, markedness perhaps does not provide a full explanation. Rather, it seemed that the participants' L1 interfered with the elements that were similar between the two languages. Vietnamese allows a nasal in a word final position. Notably, across the three tasks, expected production of clusters that contained a nasal and a voiceless stop (NSvl) ranked very high, which strongly suggests a positive transfer from L1 onto L2 final consonant production.

Considering their level of difficulty for Vietnamese ESL speakers, the 2MFCs in this study can be grouped into three groups (see Table 6), with the level of difficulty increasing from Group 1 through Group 3. It is difficult to determine, though, a cut-off level across the three groups; nevertheless, this grouping method reveals some common patterns that were also found by Benson (1988), meaning that Vietnamese ESL learners seem to have less pronunciation difficulties with SvlFvl/NSvl than with FvlFvl/NFvd, and FvlFvl/NFvd were less difficult than SvdSvd/FvdFvd/NSvd.

Table 6

Grouping of 2MFCs in Terms of Difficulty

Group 1	Group 2	Group 3
SvlFvl	SvlSvl	SvdSvd
NSvl	FvlSvl	SvdFvd
NFvl	FvlFvl	FvdSvd
	NFvd	FvdFvd
	LSvl	NSvd
	LFvl	LSvd
		LFvd

Hansen (2001) cited Greenberg (1978) that nasals can be considered to be more marked than liquids. However, this does not account for the fact that 2MCFs containing a nasal yielded higher expected articulations in this study than those containing a liquid. Again, universal markedness did not seem to hold true in this case. Given that Vietnamese does not allow liquids to end a syllable, L1 transfer was apparently one possible explanation of this result. Hansen (2001), actually, found that her Chinese informants demonstrated much less difficulty in producing final clusters with nasals, compared with those carrying liquids. She pointed out that Chinese is a language that allows nasals but not liquids in syllable codas; thus, L1 transfer seemed to be more construable than markedness.

The second question addressed in the study concerned the common modification strategies Vietnamese speakers of English used in their final consonant cluster production. Of the three identified production strategies, epenthesis was the least favored modification strategy by the participants. In this regard, Benson (1988), for example, found a high frequency of inserting a schwa after the final consonant in low proficiency Vietnamese ESL speakers. This group, however, was quite advanced; nonetheless, epenthesis was found only in one informant's production. This speaker usually inserted a schwa after an obstruent (but only in the wordlist and the interview task). According to Hansen (2001)epenthesis often occurs before a pause, so another round of re-listening to the recordings of this specific informant was done in order to pay special attention to the pauses and to double-check the pronunciation of clusters right before the pauses. There was only one true case of insertion of a schwa between the clusters ([grivəd] for *grieved*). There was no other epenthesis found and the few occurrences in the wordlist might have been due to the nature of the task—single word production where much attention is usually put on the articulation of all individual sounds. The conclusion seems to be that epenthesis was not a desirable modification strategy for 2MFCs among the advanced ESL Vietnamese learners participating in the study.

At 60% average across the three tasks, deletion appeared to be the most favored modification strategy, which is in agreement with Sato's (1984) and Osburne's (1996) findings. Elimination of the second consonant in the 2MFCs was more common than the first one. Dropping of both consonants was also found at a considerably high frequency

in the reading text and the interview which is, as Hansen (2001) argued, probably a flaw in the acquisition rather than a modification remedy. A closer look at the specific 2MFCs that were deleted revealed that they were preceded by diphthongs (e.g., /aɪ/, /eɪ/, /oʊ/, and /aʊ̃) and, in most cases, both consonants were deleted. A similar deletion was also noticed by Benson (1988) and Osburne (1996) and the explanation that was proposed pointed towards Vietnamese syllable structure, which does not usually allow consonants after diphthongs.

Changes in the features was probably the most interesting strategy to look into. Devoicing of a voiced consonant emerged as the most common strategy. Almost all 2MFCs which ended with a voiced obstruent were devoiced, whereas voicing occurred only in a few cases. In addition, as we know, aspiration of English voiceless stops /p/, /t/, /k/ does not occur in coda position, and so is the case with Vietnamese—that is, in Vietnamese, voiceless stops can be found in word-final position, where they are unreleased. However, there were instances of substitutions of /θ/ (e.g., in /naɪnθ/, which is not found in Vietnamese and is difficult for Vietnamese L2 learners of English to pronounce, with an aspirated /t^h/ (e.g., [naɪnt^h]) in a word final position, which cannot be attributed to L1 transfer.

The third question examined in the study was whether modification strategies varied across tasks. Sato (1985) found that Tarone's (1979, 1982, 1983) claim that "attention to forms" can influence speech greatly did not hold true. In Sato (1985), the participants made more mistakes in the reading task and performed better in the interview but not quite so well in the wordlist task. In this study, I found a consistent trend with the wordlist having the most desired production, followed by the text reading and the interview. The use of some sort of modification strategies was at 80.3% in the wordlist, 85.6% in the text reading, and 92.6% in the interview, which seems to suggest that pronunciation of 2MFCs may vary across task types.

In addition to the three main strategies I discussed earlier, when looking for modification strategies, I noticed that the informants appeared to frequently lengthen the vowel immediately preceding the 2MFCs, particularly in clusters containing a liquid. For instance, the word *adverb* would be pronounced as [ædvəːb] or *salt* would be pronounced as [sõːlt].̃ It should be noted that a liquid is not allowed in a word-final position in Vietnamese. Also, in some dialects of Vietnamese, there is not an equivalent of the English alveolar approximant /r/. Two of the five participants in this study spoke such dialects in which /r/ is usually substituted by /z/—for instance, /di ra/ in Vietnamese means *go out* and is pronounced in these dialects as [di za].These differences might have caused the speakers to lengthen the vowels as a compensation strategy. However, this is just a suggestion from a small set of data which needs further investigation. The second observation (which I have also touched upon earlier) is related to the frequent complete deletion of 2MFCs preceded by diphthongs, with the highest frequency being found in the text reading and the interview. This also needs further inspection in a study that specifically focuses on such structures.

Conclusion, Limitations, and Recommendations

In this study, I investigated the production of Vietnamese L2 speakers of English of English 2MFCs. The results showed that types of final clusters have a significant effect on the production output. In terms of markedness, two-member clusters of voiced obstruents were the most difficult for this group of participants. 2MFCs containing a nasal yielded the most expected productions (with the exception of nasal + voiced stop).

These findings were consistent with those of previous studies (e.g., Benson, 1988; Sato, 1984). As far as L1 transfer is concerned, given the fact that word-final consonant clusters do not exist in Vietnamese, and Vietnamese word-final consonants are not only limited in number but also carry different characteristics from those of English, evidence from this study suggested that the participants may have transferred some aspects of their L1 phonology into their acquisition of English. To be specific, 2MFCs containing a nasal led to the highest percentage of expected production. There was also a tendency of reducing the clusters to a voiceless stop (for 2MFCs of obstruents), or a single nasal (for 2MFCs containing a nasal). Furthermore, 2MFCs containing a liquid triggered a higher use of modification strategies. However, it should also be noted that about half of the clusters of this type contained the liquid /r/. Learners who had began their English learning experience following a British English pronunciation model (which was the case with all informants in the study) might have already acquired an *r*-less variety of English, which might explain the low frequency of /r/ in the participants' production of 2MFCs that contained this sound. Thus, it was difficult to conclude whether they dropped /r/ as a modification strategy or whether it was a result of a previously acquired feature. Nevertheless, /r/ occurred in many production attempts, which may suggest a developing process in the speakers' pronunciation, even at an advanced level of proficiency.

As for modification strategies, deletion and changing in features were found to be frequently employed by this group. Whereas deletion of the second consonant in the 2MFCs was the most common, there were also many cases of completely dropped 2MFCs . Deletion of both consonants occurred oftentimes with 2MFCs preceded by a diphthong, which is a vowel sound that does not take a consonant after it in the Vietnamese syllable structure.

In this study, task has shown to result in some variation in the production of final consonants in the speech of Vietnamese L2 learners of English. I have found that the interview task (where the speakers were more relaxed and less concerned with linguistic forms) resulted in the highest use of modification strategies; whereas the wordlist (where the participants had to pronounce each individual word slowly and carefully) resulted in a much lower percentage of expected articulations. This is consistent with the task model proposed by Labov (1971) and Tarone (1979, 1982, 1983), but it does not seem to confirm Sato's (1985) findings.

Despite the fact that the present study did not completely support findings of prior research, I hope to have thrown some light into the production difficulties Vietnamese advanced ESL speakers have with 2MFCs. Nevertheless, there are some limitations that the study comes with which need to be pointed out. First, the number of participants in the study was very small which has most certainly influenced the results and, respectively, has significantly limited the generalizability of the findings. Future research should aim at a much larger sample size that would allow for a quantitative analysis of the data. Secondly, I expected to have a higher inter-transcriber agreement; however, given the pressure of time and the considerate amount of transcriptions that needed to be done, the inter-transcriber agreement of .75 was an appreciable result. Inter-transcriber agreement could have been improved, though, if the informants were video recorded. Lastly, since the interview was semi-structured, this might have increased the level of formality of the task, bringing it closer to the formality level of the text reading task. This, in turn, might have influenced the participants' production and its variation across tasks.

In short, I hope this study has produced some insightful findings that contribute to the pool of findings related to ESL interlanguage phonology. The analysis presented here may prove to be beneficial not only for Vietnamese ESL teachers and learners but also ESL teachers and learners from similar language backgrounds.

Acknowledgements

I would like to thank Dr. David Bell, Dr. Scott Jarvis and Dr. Hiroyuki Oshita for giving me insightful feedback during the research process. I am indebted to Dr. Coady and Dr. Bond for their guidance and advice at the initial stage of this research. My special thanks go to Dr. Alla Zareva for her substantial revision of the paper as well as her useful suggestions for improving the quality and comprehensibility of this paper. I am most grateful to Jonathan Freeman and Kyle Hildebrandt, two graduate linguistics classmates who have helped me with one of the most crucial and time consuming procedures of this research - transcribing the data. My sincere thanks go to other faculty and friends in the Department of Linguistics for sharing experiences, knowledge, and contributing with constructive ideas during the research process. I would also like to thank my family and close friends for their patience and encouragement.

References

Anderson, J. (1987). The markedness differential hypothesis and syllable structure difficulty. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 279-291). Cambridge, MA: Newbury House.

Barley, R. (1996). Competing constraints on variation in the speech of adult Chinese learners of English. In R. Barley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 97–120). Philadelphia, PA: John Benjamin.

Benson, B. (1986). The Markedness Differential Hypothesis: Implications for Vietnamese speakers of English. In F. R. Eckman, E. A. Moravcsik, & J. R. Wirth (Eds.), *Markedness* (pp. 271–190). New York: Plenum Press.

Benson, B. (1988). Universal preference for the open syllables as an independent process in interlanguage phonology. *Language Learning, 38* (2), 221–235.

Broselow, E. (1987). An investigation of transfer in second language phonology. In G. Ioup, & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 261–274). Cambridge, MA:: Newbury House.

Celce-Murcia, M., Brinton D. M., & Goodwin J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge, MA:: Cambridge University Press.

Cutler, A., & Noris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113–121.

Eckman, F. R. (1977). Markedness and the Contrastive Analysis Hypothesis. *Language Learning, 27*(2), 315–330.

Dickerson, L., & Dickerson, W. (1977). Interlanguage phonology: Current research and future directions. In S. P. Corder & E. Roulet (Eds.), *Actes du 5e`me Colloque de Linguistique Applique´e de Neucha^tel: The notions of simplification, interlanguages, and pidgins and their relationship to second language learning* (pp. 18–29). Geneva: Droz.

Gass, S. (1980). An investigation on syntactic transfer in adult L2 learners. In R. Scarcella & S. Krashen (Eds.), *Research in second language acquisition* (pp. 132–141). Rowley, MA: Newbury House.

Gass, S., & Selinker, L. (2001). *Second language acquisition: An introductory course*. Mahwah, NJ: Laurence Erlbaum.

Greenberg, J. H. (1965). Some generalizations concerning initial and final consonant sequences. *Linguistics, 18*, 5–12.

Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant clusters. In J. H. Greenberg, (Ed.), *Universals of human language*, *Volume 2*, (pp. 243-279). Stanford: Stanford University Press.

Hansen, J. (2001). Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics, 22* (3), 338–365.

Honey, P. J. (1987). Vietnamese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 238–251). Cambridge: Cambridge University Press.

Hwa-Froelich, D., Hodson, B. W., & Edwards, H. T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech-Language Pathology, 11* (3), 264-273.

Ioup, G. & Weinberger, S. H. (1987). *Interlanguage phonology: The acquisition of a second language sound system*. Cambridge, MA: Newbury House.

Labov, W. (1971). The study of language in its social context. In J. Fishman (Ed.), *Advances in the sociology of language: Basic concepts, theories, and problems: Alternative approaches* (pp. 152–216). The Netherlands: Mouton.

Labov, W. (1989). The child as linguistic historian. Language *Variation and Change, 1*, 85–98.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Major, R. C. (1987). A model for interlanguage phonology. In G. Ioup, & S.H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 101–124). Cambridge, MA: Newbury House.

Major, R. C. (1996). Markedness in second language acquisition of consonant clusters. In R. Bayley, & D. R. Preston (Eds.), *Second language acquisition and linguistics variation* (pp. 75–96). Philadelphiap, PA: John Benjamins.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge, MA: Cambridge University Press.

Osburne, A. G. (1996). Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics, 17* (2), 164–181.

Sato, C. (1984). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning, 34*, 43–57.

Sato, C.J. (1985). Task variation in interlanguage phonology. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 181–196). Cambridge, MA: Newbury House.

Scovel, T. (1969). Foreign accents, language acquisition and cerebral dominance. Language Learning, 19, 245-254

Scovel, T. (1977). The ontogeny of the ability to recognize foreign accents. In C.C. Henning (Ed.), *Proceedings of the Second Language Research Forum* (pp. 38–47). New York: UCLA.

Tarone, E. (1979). Interlanguage as chameleon. *Language Learning, 29*, 181–191.

Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning, 32*, 69–84.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*, 142–163.

Tarone, E. (1987). The phonology of interlanguage. In G. Ioup, & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 70–87). Cambridge, MA: Newbury House.

The proficiency guidelines of the American Council for the Teaching of Foreign Languages (ACTFL). Retrieved May 9, 2008, from http://www.actfl.org/i4a/pages/index.cfm?pageid=4236

Weinberger, S. H. (1987). The influence of linguistic context on syllable simplification. In G. Ioup, & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 401–417). Cambridge, MA: Newbury House.

Wolfram, W. (1998). Tracing language history through consonant cluster reduction: *Comparative evidence from isolated dialects. Southern Journal of Linguistics, 24* (1), 17–40.

Appendix A

Wordlist

reaped	mates	laughed	reefs
ripped	mats	left	roofs
bankrupt	limits	snowdrift	play-offs
scribed	beads	grieved	leaves
scrubbed	bids	gloved	loves
backstabbed	hardheads	outlived	housewives
honk	mind	ninth	deems
hank	mend	tenth	dims
cutbank	second	midmonth	problems
salt	bold	golf	solve
silt	build	gulf	shelve
seatbelt	windshield	bookshelf	bivalve
short	garb	march	charge
shirt	curb	birch	urge
yoghurt	adverb	cornstarch	surcharge
Carl	snarl	yarn	balm
curl	squirrel	stern	film
Charles	world	whirled	palms

Appendix B

Reading text

My Childhood

I was born and raised in a rural village. Like many kids of my age, the early years were tabbed with lots of memories about the games and the songs which we played and sang together…. Most of our parents were poor farmers. Each family lived in a small cottage resembling a little shrub lumped with other cottages to form a dwarfish skirt along the sides of the river. There was no big bridge but a petite log to connect the sides, and it became a challenge for kids like us to cross the river.

Managing to produce enough food to feed the family was a difficult task; the farmers could hardly afford toys for their kids even on special occasions. Most of us suffered from malnutrition. My mom would roll my shirt up to rub and count every rib that showed. I interpreted it as a tickling game and dug my head into her bosom and giggled. She would hug me tight and her eyes would be wet with tears. I wish I had done something to soothe her pain, but my naïve mind did not let me see it through at that time.

Such was life. We village kids lived and breathed without ever having a chance to touch a cheap, much less an expensive toy. We didn't know that somewhere else on earth there existed a Toys-R-Us with its seductive, complex mechanical devices that entertained kids. Nevertheless, 'adversity brings wisdom'; we created hundreds of games. Our playthings waltzed in from whatever sources that we could find in our surroundings, simple yet inventive. We discovered enchantments in all we found and almost anything could be adopted as 'toys'. Little pebbles found at the river bank could be our playthings for days without losing their mysteriousness. The day would go fast while I learned how to hold a pebble between my fingers and shoot it at other playmates' pebbles within a blink. Sometimes we would have pebble throwing contests. He who threw the farthest would be crowned hero until someone challenged him and kicked him off the throne. I kept several pebbles of different sizes in my pant pockets all the time. My mother always complained when she sat on the flat, smooth rock in the river to wash my clothes. She would pick them out of the pants and whirl her hand high in the air to accomplish a superb throw. It always shocked me how far she could throw without even jumping up. If she had participated in our games, she would definitely have defeated us all.

Later in life, as I watched my kids playing with well polished marbles bought from Toys-R-Us shops, memories of the 'toys r'NOT us' childhood flooded in. I dashed to my room and pulled out a little box full of rough brownish pebbles. I stared at them seeing the river gradually appearing with my mother on the flat, smooth rock, picking pebbles out of my pants and throwing them away. My little playmates and I were on the bank. We whirled our hands high and jumped up into the air for a good launch. None of our pebbles ever reached the point where my mother's had landed. The river suddenly rolled up and poured into my eyes, blurring my beloved collage.

Appendix C

Map reading instrument